

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 09/980,771 B
Source: IFW116
Date Processed by STIC: 1-28-05

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 01/28/2005

PATENT APPLICATION: US/09/980,771B

TIME: 14:42:49

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\01282005\I980771B.raw

```

3 <110> APPLICANT: D'HULST, CHRISTOPHE
4     BALL, STEVEN
6 <120> TITLE OF INVENTION: STARCH GRANULES CONTAINING A RECOMBINANT POLYPEPTIDE OF
7     INTEREST, A METHOD OF OBTAINING THEM, AND THEIR USES
9 <130> FILE REFERENCE: bml-410.020
11 <140> CURRENT APPLICATION NUMBER: 09/980,771B
12 <141> CURRENT FILING DATE: 2002-01-10
14 <150> PRIOR APPLICATION NUMBER: PCT/FR00/01384
15 <151> PRIOR FILING DATE: 2000-05-19
17 <150> PRIOR APPLICATION NUMBER: FR 99/06494
18 <151> PRIOR FILING DATE: 1999-05-21
20 <160> NUMBER OF SEQ ID NOS: 15
22 <170> SOFTWARE: PatentIn Ver. 3.2
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 3117
26 <212> TYPE: DNA
27 <213> ORGANISM: Chlamydomonas reinhardtii
29 <400> SEQUENCE: 1
30 gccaggcgga gcgtatggct gttgcctcta ccagccgccc cagcagcgcg cgtcctatcg 60
31 tcatcaatgc cgcgtcggtc ggtgtcaaga agaccgcgaa ccagctgctg cgtgagcttg 120
32 ctcggtggctc cgcacgcaag tccacctcgc gctcggctgt tactgggtgcc actgggtgcc 180
33 cttgcgcgct ggacatcggt atggttgctg ctgaggtcgc cccttgggtcc aagacgggcg 240
34 gcctggggcga tgtgactggg ggctgccta ttgagctggg caagcgcggc caccgcgtca 300
35 tgaccattgc ccctcgctac gaccagtacg ctgacgcctg ggacacctcg gtggtcgtgg 360
36 acatcatggg cgagaagggt cgctacttcc actccatcaa gaaggcgctg caccgcgtgt 420
37 ggattgacca cccctgggtc ctggccaagg tctggggcaa gaccggctcc aagctgtacg 480
38 gcccccgctc cggcgctgac tacctggaca accacaagcg cttcgcccgt ttctgcaagg 540
39 ccgctattga ggctgcccgc gtgctgccct tcggcccccg cgaggactgc gtcttcgtgg 600
40 ccaacgactg gcaactccgc ctggtgcccc tctgtctgaa ggacgagtac cagcccaagg 660
41 gccagttcac caaggccaag tcggtgctgg ctatccacaa catcgccctc cagggccgca 720
42 tgtgggagga ggctttcaag gacacgaagc tgccccaggc cgcttttgac aagctggcct 780
43 tctcggaagg ctatgccaag gtttacactg aggccacccc catggaggag gacgagaagc 840
44 ccccgctgac gggaaagacc tacaagaaga tcaactggct gaagggtggc attatcgccg 900
45 ccgacaagct ggtgactgtg tcgcccact acgcgaccga gatcgctgcc gatgccgcg 960
46 gcggtgtgga gctggacacc gtcacccgc ccaagggcatt tgagggcatt gtgaacggca 1020
47 tggacattga ggagtggaa cccaagaccg acaagttcct gtctgcgccc tacgaccaga 1080
48 acagcgtcta cgccggcaag gccgcccga aggaggccct gcaggccgag ctgggcctgc 1140
49 ctgtggaccc caccgcccc ctgttcgcct tcatcgcccg cctggaggag cagaagggtg 1200
50 tggacatcat cctggccgcc ctgcccgaaga tcctggccac cccaagggtg cagatcgcca 1260
51 tcctgggtac cggcaaggcc gcctacgaga agctggtgaa cgccatcggc accaagtaca 1320
52 agggccgcgc caaggcgctg gtcaagttct cggcgccctt ggcgcacatg ctaccgcgcg 1380
53 gcgcccactt catgctggtg ccctcgcgct tcgagccctg cggcctgatc cagctgcaag 1440
54 ccatgcacta cggtagcgtg cccgtggtag cctccaccgg cggcctggtc gacaccgtca 1500

```

RAW SEQUENCE LISTING

DATE: 01/28/2005

PATENT APPLICATION: US/09/980,771B

TIME: 14:42:49

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\01282005\I980771B.raw

```

55 aggagggcggt caccggcttc cacatgggag ccctgaaccc cgacaagctg gacgaggctg 1560
56 acgcccagcgc cctggccgccc accgtgagcc gtgccagcga ggtgtttgag ggcggccgct 1620
57 accccgagat ggtggccaac tgcacagcc aggacctgtc ctgggtccaag cccgcccaga 1680
58 agtgggaggg cctgctggag gaggtggtgt acggcaaggc cggcgtggcc accgccaaga 1740
59 aggaggagat caaggtgccc gttgccgaga agatccccgg cgacctgccc gccgtgtcct 1800
60 acgcccccaa caccctgaag cccgtgtccg cctccgtgga gggcaacggc gccgcgcgc 1860
61 ccaaggtcgg caccaccgccc cccgccatgg gcgcgtggcg cgcgaccacc ccctcgggccc 1920
62 cctcgcccgcc cgccgccacc cccaaggtga ccacctacaa gcccgccttg cccgccaccg 1980
63 ccaagcccaa gaccgctggc ctcaagctgg cgggtgaggc ctccaccacc tcgacctcgg 2040
64 agaacggcgc tgccctccaac ggcaacggca acggtgcctc ggccctccaag acctcggctg 2100
65 ccaagccccc ggtctccgccc gccaccgcga agtccgccta aagcggcagt agccgcagag 2160
66 gcggcgacag catgagcggc tcgaccaaag ctgtggcagg aacggctgta gcagcggcag 2220
67 gcggccgcca ccggcgagga gcaggcttgc ggcagcagg gcgatgagct tagcggccgt 2280
68 gagcatggca ggcggaaacg tgtgtactga aatgtggtgc atgagagtgt cgtgctgtaa 2340
69 tgaagtcggt tttgcgagac cggagaaacg ccggtttggt tttgtagtgc agggcctgtg 2400
70 gtttcggttt tgcccaagtc caaaagaaga gtaacgaaac ttagtagcagta gcagagcact 2460
71 tgcgcggcgc ggcgaccacg ccggcccgtg cgcagcctgt cctgccctca gccttgtgat 2520
72 tcggcggcaa gagggcgggt ctgtacactc catccattcc aggatttttg caggctgcct 2580
73 gagagtttgc cattttgtgg gacgtgagcg gcgggacggc cgcgcggggc tctcctaccg 2640
74 cctccggcaa cggagaagtg ggaggcgtg tagcccggtg accccccaat gtagaggatg 2700
75 ggatacataa gagcgtgtgg aatggtggta aaagaggagg ggccctgggtc gcccctcgat 2760
76 ggttttgttg aggtgcagac ggcaccgtcg gcgtcaaagg cctcgcgaag gcccggtg 2820
77 cttgggctca tttttggtgc ccgtcgatga tgagagattg gccagcgggt ttttgaggct 2880
78 ggctcgaagc gaggggtttg ggaagtggag cgaggagggt tggagaaaga ggcggacatg 2940
79 cttgactgga ggtacacaaa gtggagcgtg cgacggcacg gaggcattgg cggactattg 3000
80 acccagtagt gtggaaagta gttggacctg aattctttga gaggaccgcg cattaatccg 3060
81 tgagagagta acaaagatgg cacctgaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaa 3117
84 <210> SEQ ID NO: 2
85 <211> LENGTH: 2124
86 <212> TYPE: DNA
87 <213> ORGANISM: Chlamydomonas reinhardtii
89 <220> FEATURE:
90 <223> OTHER INFORMATION: Fragment of the complete sequence of cDNA coding
91 for the GBSSI of Chlamydomonas reinhardtii
93 <220> FEATURE:
94 <221> NAME/KEY: CDS
95 <222> LOCATION: (1)..(2124)
97 <400> SEQUENCE: 2
98 atg gct gtt gcc tct acc agc cgc ccc agc agc gcg cgt cct atc gtc 48
99 Met Ala Val Ala Ser Thr Ser Arg Pro Ser Ser Ala Arg Pro Ile Val
100 1 5 10 15
102 atc aat gcc gcg tcg ttc ggt gtc aag aag acc gcg aac cag ctg ctg 96
103 Ile Asn Ala Ala Ser Phe Gly Val Lys Lys Thr Ala Asn Gln Leu Leu
104 20 25 30
106 cgt gag ctt gct cgt ggc tcc gca cgc aag tcc acc tcg cgc tcg gct 144
107 Arg Glu Leu Ala Arg Gly Ser Ala Arg Lys Ser Thr Ser Arg Ser Ala
108 35 40 45
110 gtt act ggt gcc act ggt gcc act tgc gcg ctg gac atc gtg atg gtt 192
111 Val Thr Gly Ala Thr Gly Ala Thr Cys Ala Leu Asp Ile Val Met Val

```

RAW SEQUENCE LISTING

DATE: 01/28/2005

PATENT APPLICATION: US/09/980,771B

TIME: 14:42:49

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\01282005\I980771B.raw

112	50	55	60	
114	gct gct gag gtc gcc cct tgg tcc aag acg ggc ggc ctg ggc gat gtg	240		
115	Ala Ala Glu Val Ala Pro Trp Ser Lys Thr Gly Gly Leu Gly Asp Val			
116	65 70 75 80			
118	act ggt ggc ctg cct att gag ctg gtc aag cgc ggc cac cgc gtc atg	288		
119	Thr Gly Gly Leu Pro Ile Glu Leu Val Lys Arg Gly His Arg Val Met			
120	85 90 95			
122	acc att gcc cct cgc tac gac cag tac gct gac gcc tgg gac acc tcg	336		
123	Thr Ile Ala Pro Arg Tyr Asp Gln Tyr Ala Asp Ala Trp Asp Thr Ser			
124	100 105 110			
126	gtg gtc gtg gac atc atg ggc gag aag gtc cgc tac ttc cac tcc atc	384		
127	Val Val Val Asp Ile Met Gly Glu Lys Val Arg Tyr Phe His Ser Ile			
128	115 120 125			
130	aag aag ggc gtg cac cgc gtg tgg att gac cac ccc tgg ttc ctg gcc	432		
131	Lys Lys Gly Val His Arg Val Trp Ile Asp His Pro Trp Phe Leu Ala			
132	130 135 140			
134	aag gtc tgg ggc aag acc ggc tcc aag ctg tac ggc ccc cgc tcc ggc	480		
135	Lys Val Trp Gly Lys Thr Gly Ser Lys Leu Tyr Gly Pro Arg Ser Gly			
136	145 150 155 160			
138	gct gac tac ctg gac aac cac aag cgc ttc gcc ctg ttc tgc aag gcc	528		
139	Ala Asp Tyr Leu Asp Asn His Lys Arg Phe Ala Leu Phe Cys Lys Ala			
140	165 170 175			
142	gct att gag gct gcc cgc gtg ctg ccc ttc ggc ccc ggc gag gac tgc	576		
143	Ala Ile Glu Ala Ala Arg Val Leu Pro Phe Gly Pro Gly Glu Asp Cys			
144	180 185 190			
146	gtc ttc gtg gcc aac gac tgg cac tcc gcc ctg gtg ccc gtc ctg ctg	624		
147	Val Phe Val Ala Asn Asp Trp His Ser Ala Leu Val Pro Val Leu Leu			
148	195 200 205			
150	aag gac gag tac cag ccc aag ggc cag ttc acc aag gcc aag tcg gtg	672		
151	Lys Asp Glu Tyr Gln Pro Lys Gly Gln Phe Thr Lys Ala Lys Ser Val			
152	210 215 220			
154	ctg gct atc cac aac atc gcc ttc cag ggc cgc atg tgg gag gag gct	720		
155	Leu Ala Ile His Asn Ile Ala Phe Gln Gly Arg Met Trp Glu Glu Ala			
156	225 230 235 240			
158	ttc aag gac acg aag ctg ccc cag gcc gcc ttt gac aag ctg gcc ttc	768		
159	Phe Lys Asp Thr Lys Leu Pro Gln Ala Ala Phe Asp Lys Leu Ala Phe			
160	245 250 255			
162	tcg gac ggc tat gcc aag gtt tac act gag gcc acc ccc atg gag gag	816		
163	Ser Asp Gly Tyr Ala Lys Val Tyr Thr Glu Ala Thr Pro Met Glu Glu			
164	260 265 270			
166	gac gag aag ccc ccg ctg acg gga aag acc tac aag aag atc aac tgg	864		
167	Asp Glu Lys Pro Pro Leu Thr Gly Lys Thr Tyr Lys Lys Ile Asn Trp			
168	275 280 285			
170	ctg aag ggt ggc att atc gcc gcc gac aag ctg gtg act gtg tcg ccc	912		
171	Leu Lys Gly Gly Ile Ile Ala Ala Asp Lys Leu Val Thr Val Ser Pro			
172	290 295 300			
174	aac tac gcg acc gag atc gct gcc gat gcc gcc ggc ggt gtg gag ctg	960		
175	Asn Tyr Ala Thr Glu Ile Ala Ala Asp Ala Ala Gly Gly Val Glu Leu			
176	305 310 315 320			

RAW SEQUENCE LISTING

DATE: 01/28/2005

PATENT APPLICATION: US/09/980,771B

TIME: 14:42:49

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\01282005\I980771B.raw

178	gac acc gtc atc cgc gcc aag ggc att gag ggc att gtg aac ggc atg	1008
179	Asp Thr Val Ile Arg Ala Lys Gly Ile Glu Gly Ile Val Asn Gly Met	
180	325 330 335	
182	gac att gag gag tgg aac ccc aag acc gac aag ttc ctg tct gcg ccc	1056
183	Asp Ile Glu Glu Trp Asn Pro Lys Thr Asp Lys Phe Leu Ser Ala Pro	
184	340 345 350	
186	tac gac cag aac agc gtc tac gcc ggc aag gcc gcc gcc aag gag gcc	1104
187	Tyr Asp Gln Asn Ser Val Tyr Ala Gly Lys Ala Ala Ala Lys Glu Ala	
188	355 360 365	
190	ctg cag gcc gag ctg ggc ctg cct gtg gac ccc acc gcc ccc ctg ttc	1152
191	Leu Gln Ala Glu Leu Gly Leu Pro Val Asp Pro Thr Ala Pro Leu Phe	
192	370 375 380	
194	gcc ttc atc ggc cgc ctg gag gag cag aag ggt gtg gac atc atc ctg	1200
195	Ala Phe Ile Gly Arg Leu Glu Glu Gln Lys Gly Val Asp Ile Ile Leu	
196	385 390 395 400	
198	gcc gcc ctg ccc aag atc ctg gcc acc ccc aag gtg cag atc gcc atc	1248
199	Ala Ala Leu Pro Lys Ile Leu Ala Thr Pro Lys Val Gln Ile Ala Ile	
200	405 410 415	
202	ctg ggt acc ggc aag gcc gcc tac gag aag ctg gtg aac gcc atc ggc	1296
203	Leu Gly Thr Gly Lys Ala Ala Tyr Glu Lys Leu Val Asn Ala Ile Gly	
204	420 425 430	
206	acc aag tac aag ggc cgc gcc aag ggc gtg gtc aag ttc tcg gcg ccc	1344
207	Thr Lys Tyr Lys Gly Arg Ala Lys Gly Val Val Lys Phe Ser Ala Pro	
208	435 440 445	
210	ctg gcg cac atg ctc acc gcc ggc gcc gac ttc atg ctg gtg ccc tcg	1392
211	Leu Ala His Met Leu Thr Ala Gly Ala Asp Phe Met Leu Val Pro Ser	
212	450 455 460	
214	cgc ttc gag ccc tgc ggc ctg atc cag ctg cac gcc atg cac tac ggt	1440
215	Arg Phe Glu Pro Cys Gly Leu Ile Gln Leu His Ala Met His Tyr Gly	
216	465 470 475 480	
218	acc gtg ccc gtg gta gcc tcc acc ggc ggc ctg gtc gac acc gtc aag	1488
219	Thr Val Pro Val Val Ala Ser Thr Gly Gly Leu Val Asp Thr Val Lys	
220	485 490 495	
222	gag ggc gtc acc ggc ttc cac atg ggc gcc ctg aac ccc gac aag ctg	1536
223	Glu Gly Val Thr Gly Phe His Met Gly Ala Leu Asn Pro Asp Lys Leu	
224	500 505 510	
226	gac gag gct gac gcc gac gcc ctg gcc gcc acc gtg cgc cgt gcc agc	1584
227	Asp Glu Ala Asp Ala Asp Ala Leu Ala Ala Thr Val Arg Arg Ala Ser	
228	515 520 525	
230	gag gtg ttt gcg ggc ggc cgc tac ccc gag atg gtg gcc aac tgc atc	1632
231	Glu Val Phe Ala Gly Gly Arg Tyr Pro Glu Met Val Ala Asn Cys Ile	
232	530 535 540	
234	agc cag gac ctg tcc tgg tcc aag ccc gcc cag aag tgg gag ggc ctg	1680
235	Ser Gln Asp Leu Ser Trp Ser Lys Pro Ala Gln Lys Trp Glu Gly Leu	
236	545 550 555 560	
238	ctg gag gag gtg gtg tac ggc aag ggc ggc gtg gcc acc gcc aag aag	1728
239	Leu Glu Glu Val Val Tyr Gly Lys Gly Gly Val Ala Thr Ala Lys Lys	
240	565 570 575	
242	gag gag atc aag gtg ccc gtt gcc gag aag atc ccc ggc gac ctg ccc	1776

RAW SEQUENCE LISTING

DATE: 01/28/2005

PATENT APPLICATION: US/09/980,771B

TIME: 14:42:49

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\01282005\I980771B.raw

```

243 Glu Glu Ile Lys Val Pro Val Ala Glu Lys Ile Pro Gly Asp Leu Pro
244      580      585      590
246 gcc gtg tcc tac gcc ccc aac acc ctg aag ccc gtg tcc gcc tcc gtg 1824
247 Ala Val Ser Tyr Ala Pro Asn Thr Leu Lys Pro Val Ser Ala Ser Val
248      595      600      605
250 gag ggc aac ggc gcc gcc gcg ccc aag gtc ggc acc acc gcc ccc gcc 1872
251 Glu Gly Asn Gly Ala Ala Ala Pro Lys Val Gly Thr Thr Ala Pro Ala
252      610      615      620
254 atg ggc gcg tgg cgc gcg acc acc ccc tcg ggc ccc tcg ccc gcc gcc 1920
255 Met Gly Ala Trp Arg Ala Thr Thr Pro Ser Gly Pro Ser Pro Ala Ala
256 625      630      635      640
258 gcc acc ccc aag gtg acc acc tac aag ccc gcc ctg ccc gcc acc gcc 1968
259 Ala Thr Pro Lys Val Thr Thr Tyr Lys Pro Ala Leu Pro Ala Thr Ala
260      645      650      655
262 aag ccc aag acc gct ggc ctc aag ctg gcc ggt gag gcc tcc acc acc 2016
263 Lys Pro Lys Thr Ala Gly Leu Lys Leu Ala Gly Glu Ala Ser Thr Thr
264      660      665      670
266 tcg acc tcg gag aac ggc gct gcc tcc aac ggc aac ggc aac ggt gcc 2064
267 Ser Thr Ser Glu Asn Gly Ala Ala Ser Asn Gly Asn Gly Asn Gly Ala
268      675      680      685
270 tcg gcc tcc aag acc tcg gct gcc aag ccc ctg gtc tcc gcc gcc acc 2112
271 Ser Ala Ser Lys Thr Ser Ala Ala Lys Pro Leu Val Ser Ala Ala Thr
272      690      695      700
274 cgc aag tcc gcc 2124
275 Arg Lys Ser Ala
276 705
281 <210> SEQ ID NO: 3
282 <211> LENGTH: 708
283 <212> TYPE: PRT
284 <213> ORGANISM: Chlamydomonas reinhardtii
286 <400> SEQUENCE: 3
287 Met Ala Val Ala Ser Thr Ser Arg Pro Ser Ser Ala Arg Pro Ile Val
288 1 5 10 15
290 Ile Asn Ala Ala Ser Phe Gly Val Lys Lys Thr Ala Asn Gln Leu Leu
291 20 25 30
293 Arg Glu Leu Ala Arg Gly Ser Ala Arg Lys Ser Thr Ser Arg Ser Ala
294 35 40 45
296 Val Thr Gly Ala Thr Gly Ala Thr Cys Ala Leu Asp Ile Val Met Val
297 50 55 60
299 Ala Ala Glu Val Ala Pro Trp Ser Lys Thr Gly Gly Leu Gly Asp Val
300 65 70 75 80
302 Thr Gly Gly Leu Pro Ile Glu Leu Val Lys Arg Gly His Arg Val Met
303 85 90 95
305 Thr Ile Ala Pro Arg Tyr Asp Gln Tyr Ala Asp Ala Trp Asp Thr Ser
306 100 105 110
308 Val Val Val Asp Ile Met Gly Glu Lys Val Arg Tyr Phe His Ser Ile
309 115 120 125
311 Lys Lys Gly Val His Arg Val Trp Ile Asp His Pro Trp Phe Leu Ala
312 130 135 140

```

VERIFICATION SUMMARY

DATE: 01/28/2005

PATENT APPLICATION: **US/09/980,771B**

TIME: 14:42:50

Input Set : **A:\PTO.FG.txt**

Output Set: **N:\CRF4\01282005\I980771B.raw**